Experimental article – A balance between image quality and effective dose in orbital X-ray screening for ferromagnetic IOFBs: a pilot study

Gabrielle Hart[a], Sarah Jessop[a], Ana Rita Santiago[b], Abbas Samara[c], Benedicte Markali[d], Yann Cottier[e], Joana Guerreiro[b], Erik Normann Andersen[d], H. Momoniat[a]*, José Jorge[e], Andrew England[a]

a) *School of Health Sciences, University of Salford, Manchester, United Kingdom*
b) *Lisbon School of Health Technology (ESTeSL), Polytechnic Institute of Lisbon, Portugal*
c) *Department of Medical Imaging and Radiation Therapy, Hanze University of Applied Sciences, Groningen, The Netherlands*
d) *Department of Life Sciences and Health, Radiography, Oslo and Akerhus University College of Applied Science, Oslo, Norway*
e) *Haute École de Santé Vaud – Filière TRM, University of Applied Sciences and Arts of Western Switzerland, Lausanne, Switzerland*

* *Acknowledgments to College of Radiographers Industry Partnership Scheme (CoRIPS) for the grant awarded to H. Momoniat.*

ABSTRACT

Purpose: To investigate whether standard X-ray acquisition factors for orbital radiographs are suitable for the detection of ferromagnetic intra-ocular foreign bodies in patients undergoing MRI.
Method: 35 observers, at varied levels of education in radiography, attending a European Dose Optimisation EURASMUS Summer School were asked to score 24 images of varying acquisition factors against a clinical standard (reference image) using two alternative forced choice. The observers were provided with 12 questions and a 5 point Likert scale. Statistical tests were used to validate the scale, and scale reliability was also measured. The images which scored equal to, or better than, the reference image (36) were ranked alongside their corresponding effective dose (E), the image with the lowest dose equal to or better than the reference is considered the new optimum acquisition factors.
Results: Four images emerged as equal to, or better than, the reference in terms of image quality. The images were then ranked in order of E. Only one image that scored the same as the reference had a lower dose. The reference image had a mean E of 3.31μSv, the image that scored the same had an E of 1.8μSv.
Conclusion: Against the current clinical standard exposure factors of 70kVp, 20mAs and the use of an anti- scatter grid, one image proved to have a lower E whilst maintaining the same level of image quality and lesion visibility. It is suggested that the new exposure factors should be 60kVp, 20mAs and still include the use of an anti-scatter grid.

INTRODUCTION

A case from the 1980's, highlighted by Kelly et al, saw an American man being blinded by an undetected metal fragment when undergoing a Magnetic Resonance Imaging (MRI) scan. Even though he provided a history of Intra Orbital Foreign Body (IOFB) to the radiographers and underwent a subsequent plain X-ray examination, the fragment was undetected upon first review of the image[1]. After the MRI incident the IOFB was seen on the image, suggesting that the technique used was not optimised and the quality of the image was so low that human error meant severe harm to the patient, highlighting the importance of image optimization while maintaining As Low As Reasonably Achievable (ALARA)[2] principle.

Prior to MRI scan, a safety questionnaire is a good instrument to evaluate whether a patient is at "high risk" of having an IOFB and therefore an orbit X-ray candidate[3]. Although, there is a case which the patient denied having any IOFB and later he developed hyphema due to a ferromagnetic fragment in the eye[4].

The lens of the eye is considered to be one of the most radiosensitive tissues of the human body and high or repeated direct exposure causes lens clouding or cataracts, a type of visual impairment[5]. For that reason it is of paramount importance to optimise dose when performing an orbit X-ray.

This study will investigate image quality and dose optimi-

sation in Computed Radiography (CR) in relation to orbital X-rays for MRI screening.

METHODS AND MATERIALS

Equipment and phantom setup

An adult anthropomorphic head phantom was positioned for a postero-anterior (PA) projection of the orbits in accordance with standard radiographic texts[6-7] (Figure 1). Images were acquired using a Wolverson Acroma X-ray unit (high frequency generator with VARIAN 130 HS standard X-ray tube with a total filtration of 3mm Aluminium equivalent). The source-to-image receptor distance (SID) was set at 100cm and all images were acquired using the same 18 x 24cm CR image receptor (IR). The primary X-ray beam was collimated to include the lateral skull margins and the whole orbital region and was thus fixed at 21.5 x 8.5cm. An Agfa 35-X digitizer (Agfa-Gavaert Corp, Mortsel, Belgium) was used to process the images using a skull look up table.

IOFB simulation

Five ferromagnetic IOFBs (<1.0mm) were fixed to the anterior aspect of the orbital region of the phantom on the right eye in a pre-determined distribution (Figure 2). The left eye was maintained free from IOFB and would be used to simulate a normal examination.

Image acquisition

A set of images, for the purpose of both image quality and dosemetric analyses, were generated using the phantom and the following acquisition parameters. For peak tube potential, images were acquired at 10kV increments from 60 to 90kVp. For mAs, 5.0, 20.0 and 40.0 were selected. For the first set of images the IR was placed in the vertical bucky which included a secondary radiation grid (ratio 10:1, 40 lines/cm). A second set of images was acquired without a radiation grid using the same kVp and mAs settings.

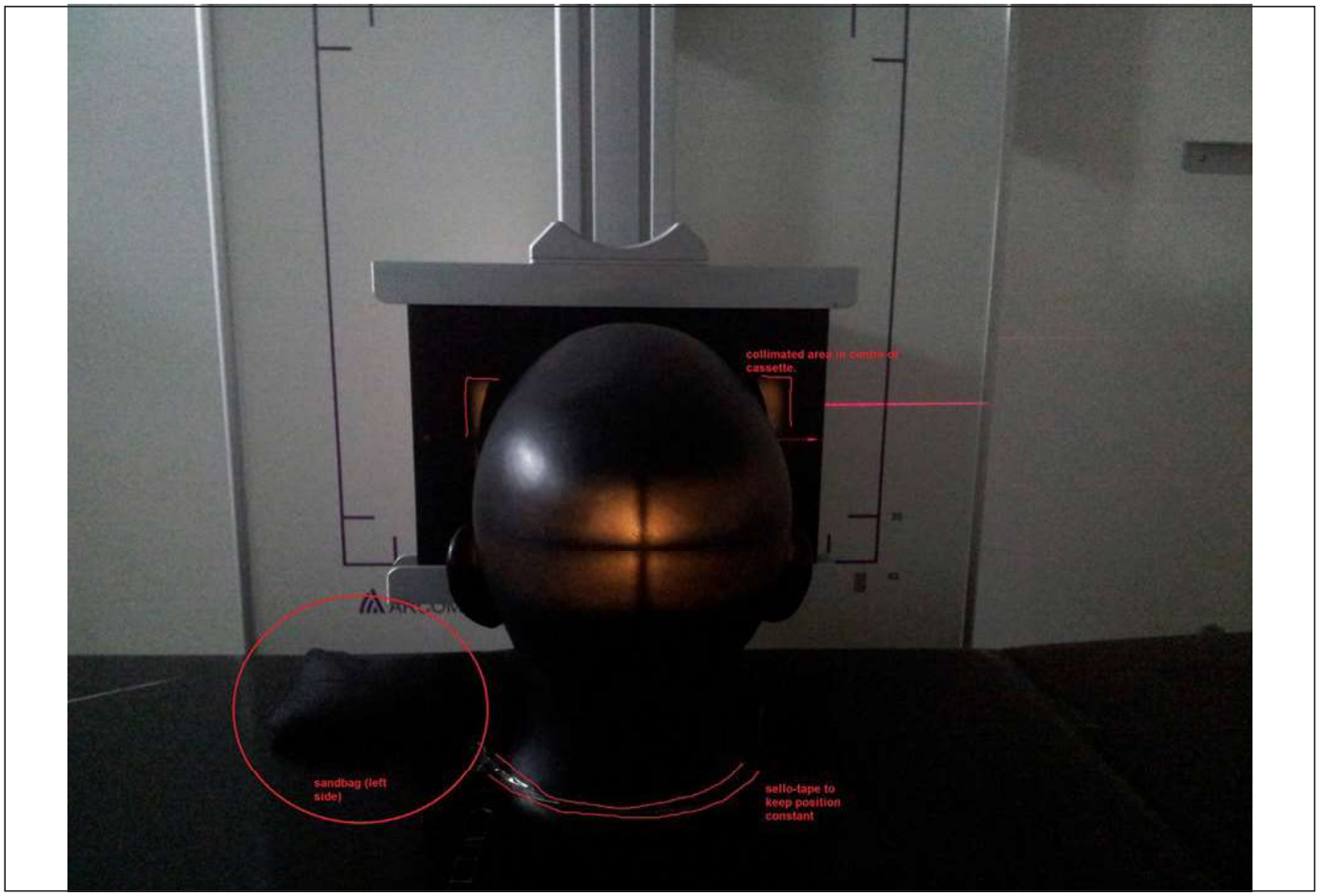

Figure 1: *An illustration of the X-ray equipment and phantom setup used in this study. The annotations represent the collimation, the sandbag which steadied the phantom head (to the bottom left) and sellotape used to ensure no movement.*

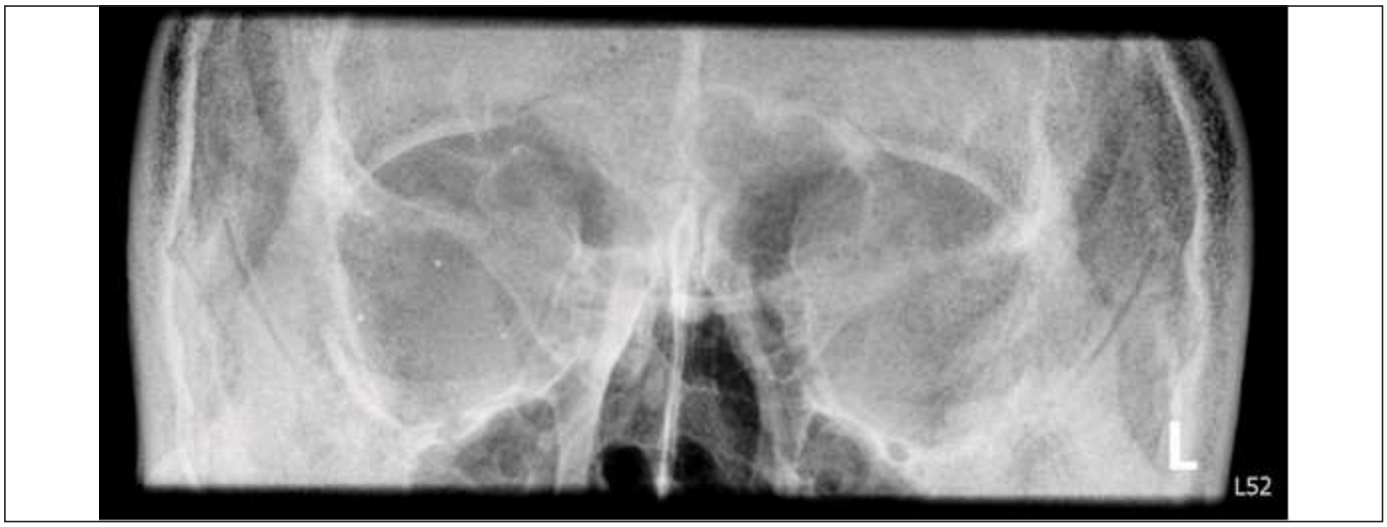

Figure 2: The image sample demonstrates a normal eye (left) and one eye (right) with 5 IOFBs.

Figure 3: A PA orbital radiograph demonstrated the location and size of the two ROIs used in the physical measurement of image quality.

A total of 24 different acquisition factor combinations were selected and acquired. For each of the settings, three X-ray exposures were obtained and the Dose-Area-Product (DAP) values were recorded. At each acquisition parameter combination a single image was send to an archive and the Exposure Index (LgM) was recorded.

Image quality analysis

Physical measures

Acquired images were first evaluated using physical measures of image quality, to validate the image quality scale and gave an objective measure of image quality. Mean and standard deviation pixel value at two locations were calculated using the ImageJ software (National Institute of Health, Bethesda, MD) using a fixed sized region of interest (ROI). Two ROIs (S1 and S2) were plotted (Figure 3) and from this signal-to-noise (SNR) and contrast-to-noise (CNR) values were calculated. SNR was defined as the mean pixel value divided by the standard deviation for each ROI, CNR was defined as the difference between the mean pixel values divided by the standard deviation between each ROI. These methodologies have been used in similar experiments[8-9].

Perceptual (visual) tests

35 observers from the Netherlands, Switzerland, Portugal, Norway and UK volunteered for the image quality test (mean age = 26.1, range = 19 - 56). All observers had normal to corrected-to-normal vision, although, one participant who would usually wear glasses had forgotten them. The scale was produced through literature review and focus group discussion. Reliability and validation were tested. This approach has been

used in similar radiographic projects reported in the literature[8,10]. Observers were radiographers (students or qualified practitioners) on a European Dose Optimisation EURASMUS Summer School. Images were initially analysed visually used two alternative forced choice comparisons (2AFC)[8]. 2AFC assesses the psychometric responses of observers who are presented with two separate images and has been used extensively within radiography to compare image quality[8,11-14]. Limited resources meant 2AFC was as follows, two observers shared one screen and the set up was modified as follows; on the top of the screen, two reference images were fixed, on the bottom the remaining images were presented to each observer in a random order. Selection of a reference image was based on those parameters which reflect typical clinical averages, this was decided by discussions between the study researchers (70kVp, 20mAs and inclusion of an anti-scatter radiation grid). For each image, observers were required to indicate their level of agreement for each scale item against the reference image, where 1 was much worse, 2 worse, 3 the same, 4 better and 5 much better (Table 1). A score of 3 indicated a comparable image to the reference image for that specific criterion.

Table 1: Summary of the perceptual image quality scoring questionnaire (scale) used in the experiment

Contrast between air-filled structures and the surrounding tissues/ structures	
Trabecular pattern of the visualised bones	
Sharpness of the orbital rim	
Visibility of the superior orbital fissure	
Quality of noise	
With respect to the visualised lesions:	Brightness
	Contrast
	Visibility
The scale consisted of a total of 12 items.	

Test procedure

Two participants at a time viewed the reference and comparison images on a split screen 30 inch Eizo MX300 (Eizo Corp, Hakusan, Ishikawa, Japan) liquid crystal display (LCD) monitor with a resolution of 2 megapixels, as stated above. Monitors were calibrated to DICOM greyscale standard display function (GSDF) and the ambient lighting conditions were kept constant and dimmed (i.e., 32 Lux) in accordance with the European Guidelines on Quality Criteria for Diagnostic Radiographic Images[15]. Noise levels and interruptions to image review were minimised using a sign on the door. Full instructions to observers were given at the start of the visual assessments and observers also were subject to a short training session prior. Definitions for each image quality criterion were provided in writing together with an anatomy and IOFB location visual aid (Appendix A).

Scale validation

Testing of the scale included the use of both physical measures and scale questionnaires returned from the first 16 participants. Correlations between SNR and mean image quality scores (total per image) have been used previously[8]. Using all data collected in our study, there was almost no correlation between total image quality score and SNR (S_1 R^2 = 0.022, p=0.910, S_2 R^2 =0.031, p=0.886; Figure 4).

For CNR there was a moderate positive correlation R^2 = 0.302, p<0.005 (Figure 4) against total score.

Validating a scale which includes both normal anatomy and simulated lesions is likely to require metrics other than SNR and CNR. Evidence presented above confirms that image quality scores do have some relationship with SNR and CNR. Time constraints only allowed for one test, re-test. The ICC was 0.508 (95% CI). Rosner (2011) suggested that values in the region of 0.40-0.75 indicate fair to good reproducibility.

Based on a review of SNR and mean image quality scores (IQS) from 35 participants there were still no significant correlations identified with respect to the full image quality scale (S_1 : R^2 = 0.001, p = 0.884, S_2 : R^2 = 0.009, p = 0.655).

There was statistically significant correlation between SNR and the average IQS for question 5 (S_1 : R^2 = 0.595, p < 0.001, S_2 : R^2 = 0.588, p < 0.001) (Figure 5).

Further validation analyses were undertaken on a per question basis. CNR did demonstrate a moderate positive correlation with mean IQS for question 1 (R^2 = 0.446, p <0.001), question 7 (R^2 = 0.449, p <0.001), question 8 R^2 = 0.432, p<0.001), question 10 (R^2 = 0.413, p = 0.001), and question 12 (R^2 = 0.401, p = 0.001). CNR demonstrated a lower positive correlation with mean IQS for question 9 (R^2 = 0.338, p = 0.003), question 11 (R^2 = 0.374, p = 0.002) and for the total IQS (R^2 = 0.380, p = 0.001).

Evidence presented above and in the early stage (n=16) scale validation indicates that IQS do have some relationship with SNR and CNR.

In order to test the reliability of the image scoring system inter-observer variability ICC values were calculated for each image.

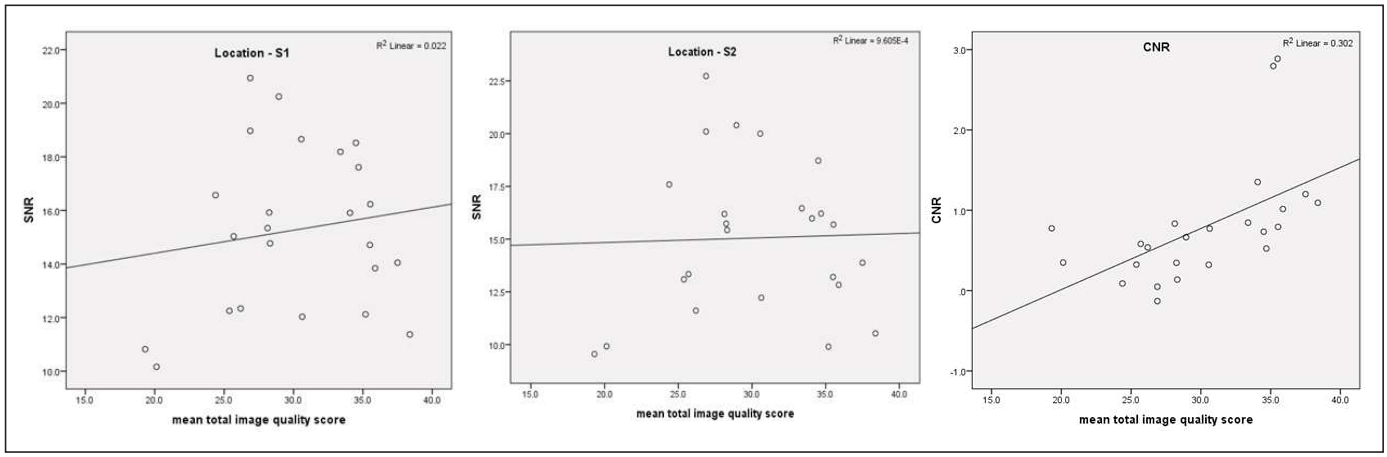

Figure 4: Scatter plots of SNR and CNR when compared to mean total image quality scores of question 5.

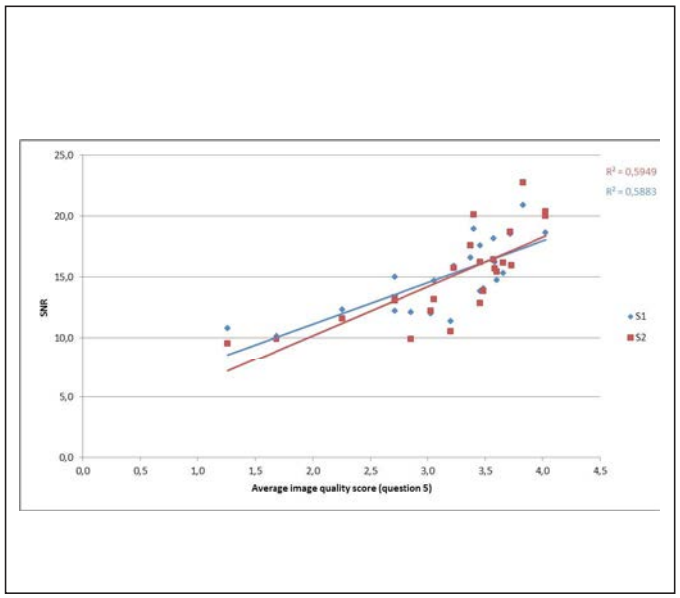

Figure 5: Scatter graphs of SNR compared to average total image quality scores of question 5.

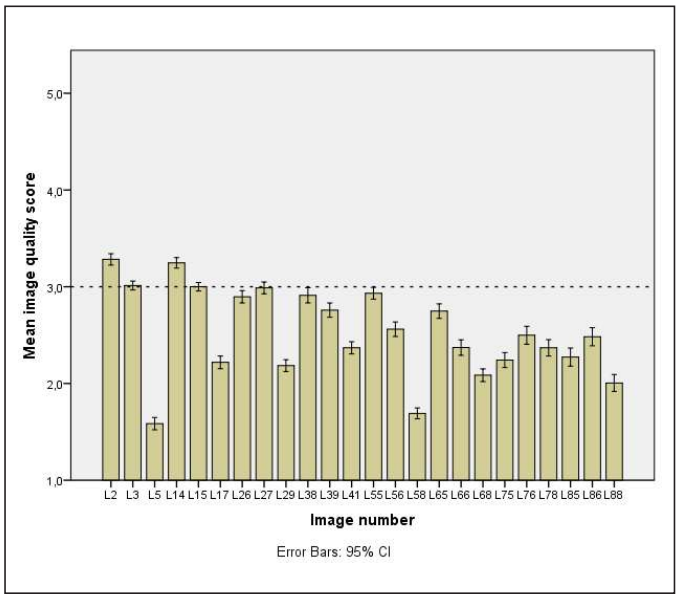

Figure 6. The histogram showing all mean image quality scores for all 24 images.

This used a 2-way mixed effect model for absolute agreement and the SPSS computer software (IBM Corp, 2011). The ICC (N = 35) is 0.456 (95% CI). When interpreting ICC, Rosner, suggested that values in the region of 0.40-0.75 indicate fair to good[16].

Radiation dosimetry

DAP readings were recorded during acquisition. An average of three readings was taken for each image acquisition. Effective dose (E) were calculated from the DAP using Monte Carlo simulation software (PCXMC 2.0). The PCXMC, Monte Carlo base computer software uses computational hermaphrodite phantom defined by mathematical expressions to compute organ and E of patients of different ages and sizes in freely adjustable X-ray projections and other examination conditions used in radiology[17]. PCXMC calculates Es using ICRP, 2007 publication 103 recommendations[15,18]. The reliability of this software is supported by literature demonstrating results in close agreement with dose measurements and calculations of other phantom models

Statistical analysis

All IQS were transferred to SPSS (IBM Corp., 2011) and mean scores across 11 criteria were calculated, due to an understanding that many observers did not understand question 6. In terms of dose optimization images close to reference IQS (mean L15 = 3.0) were identified. Identified images (4 images) were compared with the reference image by non-parametric Wilcoxon matched-pair signed-rank test (corrected for multiple comparisons).

RESULTS

Perceptual image quality

Figure 6 shows the mean IQS for each of image with a range from 1.584 (St. Dev = 0.456) to 3.283 (St. Dev = 0.340). The mean values which scored above the reference image (represented by the dotted line) suggest better IQS. Several images scored just below the reference image.

Images for further analysis were identified by their mean IQS (compared with the reference image) and E (μSv). Wilcoxon matched-pair signed-rank test showed the mean IQS for images L3 (P= 0.963), L27 (P = 0.945) and L55 (P = 0.803) were not statistically significant from the reference image L15. However, image L2 (P = .000) was statistically significant compared to the reference image, L15, the Wilcoxon test is not able to differentiate which direction the mean difference is in. But the difference in the mean IQS for L2 and L15 (Table 2) suggests that observers rate L2 significantly higher than the reference image L15.

Table 2: Describes the descriptive values for each tested image and p

Image Name	kVp	mAs	E (μSv)	Mean IQ score	Std. Dev score	Wilcoxon signed rank (with ref image L15)
L3	60	20	1.821	3.01	0.31	p > 0.05
L15 (ref img)	70	20	3.308667	3	0.27	-
L55	60	40	3.531333	2.93	0.31	p > 0.05
L2	60	40	3.762667	3.28	0.34	p < 0.05
L27	80	20	5.025	2.99	0.33	p > 0.05

DISCUSSION

Study findings

The results from this pilot study suggest that using 60kVp 20mAs does not significantly affect the perceived image quality when compared with the clinical average which is 70kVp 20mAs. However the 60kVp 20mAs (1.821μSv) reduce the E with 45% compared with the reference image (3.308μSv). Some of the images in the Figure 6 scored slightly higher, lower, or close to the reference image but were excluded from further study based upon their E. One of the images (6.279μSv) that was excluded had an 89% increase in E compared with the reference image but scored higher. The perceived image quality of image L2 (60kVp 40mAs, 3.762μSv) is significantly higher than the reference image but it provides a 13% dose increase when compared with the reference image (see Table 2).

Literature comparison

The results show that, as with kVp decrease, E decreases but IQS remain very similar. This is supported in the work of Allen et al, whose research states that a 10kVp decrease will see a decrease in E with no real compromise in image quality[19]. This is supported in the above results as the reference image has an E of 3.30μSv, and where the mAs stays the same and the kVp decreases from 70 to 60, the E decreases to 1.82μSv. The mean IQS for both images is very similar, with a small difference of 0.1 in favour of the lower kVp image.

Ma et al also agree that the image quality remains the same while decreasing E between 70kVp and 60kVp. They see similar results in their study where the dose for an acquisition at 70kVp is around 3μSv, and when it is reduced to 60kVp, the dose is reduced to around 2μSv11. This reflects the above results.

Implications on clinical practice

After more in depth research is conducted, presuming the results are similar to the above, implications on clinical practice may be that the new, lower acquisition factors are trialed in only a and the image quality tested by experienced and qualified film readers to see whether they can still see any IOFBs with the lower exposure. If the film readers still maintain a high rate of IOFB identification then the new exposure may become the standard.

Recommendations for improvement

Several factors may have influenced the study, subsequently limiting it. The first was related to the images for analysis. Problems occurred when the observers noticed differences in shuttering throughout the images, which occurred due to a post processing error. This meant that observers found it more difficult to compare the images fully, and the investigators found it harder to place the ROIs. Some observers complained that the LCD screens had a coloured tint and that changed their perception to some degree, although this was an uncommon report. Some observers reported a misunderstanding of question 6, these results were subsequently removed for all observers.

The pre-questionnaire observers were asked to fill out before performing image analysis highlighted the variety in the participants. This meant level of experience within the participants could be monitored. A range of people at different levels in their radiographic education (whether qualified or student) were asked to participate. Students at a lower level may have been less experienced in image evaluation, but this was controlled as much as possible by universal training. Only two participants highlighted this as a problem and subsequently withdrew from the study voluntarily. The experience level could have affected the ICC but we can't discount other variables.

The conditions in the room were controlled as much as possible; however other groups of researchers were using it. This meant that some noise (talking) and light (from the door opening and closing) were exposed to the participants while they graded the images. This may have been distracting but was minimised and was not reported as a problem.

Recommendations for further work

In further studies, the participants asked could be controlled, and invite only qualified radiographers alongside reporting radiographers and radiologists to grade the images. This change may improve the external validity of the findings due to the increase in relevant experience.

Different projections could be acquired to try and minimise dose such as a caudal angle as suggested by Bontrager et al[20].

It may be interesting to repeat the investigation using a Direct Radiography (DR) system. The reduction in exposure from film to CR was drastic (75kVp and 40mAs with a distance of 90cm to 70kVp, 20mAs and a distance of 100cm) and so the E decreased largely, it is likely the dose would decrease with the progression of technology.

CONCLUSION

The results of the study indicate that there is an opportunity in CR radiography to decrease the acquisition factors, namely kVp, in orbital X-rays. The radiograph that demonstrated 60kVp, 20mAs and 100cm SID was rated similarly in image quality to the reference, or clinical average, and provides a dose of 1.8μSv rather than the clinical average of 3.3μSv.

REFERENCES

1. Kelly WM, Paglen PG, Pearson JA, San Diego AG, Solomon MA. Ferromagnetism of intraocular foreign body causes unilateral blindness after MR study. AJNR Am J Neuroradiol. 1986;7(2):243-5.

2. Kohn M, Moores B, Schibilla H, Schneider K, Stender H, Stieve F, et al. European guidelines on quality criteria for diagnostic radiographic images in paediatrics. Luxembourg: Office for Official Publications of the European Communities; 1996.

3. Bailey W, Robinson L. Screening for intra-orbital metallic foreign bodies prior to MRI: review of the evidence. Radiography. 2007;13(1):72-80.

4. Ta CN, Bowman RW. Hyphema caused by a metallic intraocular foreign body during magnetic resonance imaging. Am J Ophthalmol. 2000;129(4):533-4.

5. Chodick G, Bekiroglu N, Hauptmann M, Alexander BH, Freedman DM, Doody MM, et al. Risk of cataract after exposure to low doses of ionizing radiation: a 20-year prospective cohort study among US radiologic technologists. Am J Epidemiol. 2008;168(6):620-31.

6. Whitley AS, Sloane C, Hoadley G, Moore AD. Clark's positioning in radiography. 12th ed. London: CRC Press; 2005.

7. Ballinger PW, Frank ED. Merrill's atlas of radiographic positions & radiologic procedures. 10th ed. London: Mosby; 2002.

8. Mraity H, England A, Akhtar I, Aslam A, De Lange R, Momoniat H, et al. Development and validation of a psychometric scale for assessing PA chest image quality: a pilot study. Radiography. 2014;20(4):312-7.

9. Lin Y, Luo H, Dobbins JT, Page McAdams H, Wang X, Sehnert WJ, et al. An image-based technique to assess the perceptual quality of clinical chest radiographs. Med Phys. 2012;39(11):7019-31.

10. Mraity H, England A, Hogg P. Developing and validating a psychometric scale for image quality assessment. Radiography. 2014;20(4):306-11.

11. Ma WK, Hogg P, Tootell A, Manning D, Thomas N, Kane T, et al. Anthropomorphic chest phantom imaging: the potential for dose creep in computed radiography. Radiography.

2013;19(3):207-11.

12. Fechner GT. Elemente der psychophysik. 2nd ed. Leipzig: Breitkopf und Härtel; 1889.

13. Lança L, Franco L, Ahmed A, Harderwijk M, Marti C, Nasir S, et al. 10 kVp rule – An anthropomorphic pelvis phantom imaging study using a CR system: impact on image quality and effective dose using AEC and manual mode. Radiography. 2014;20(4):333-8.

14. Reis C, Gonçalves J, Klompmaker C, Bárbara AR, Bloor C, Hegarty R, et al. Image quality and dose analysis for a PA chest X-ray: comparison between AEC mode acquisition and manual mode using the 10 kVp 'rule'. Radiography. 2014;20(4):339-45.

15. ICRP. The 2007 recommendations of the International Commission on Radiological Protection [Internet]. Ottawa: ICRP; 2007 [cited 2014 Aug 12]. Available from: http://www.icrp.org/publication.asp?id=ICRP%20Publication%20103

16. Rosner B. Fundamentals of biostatistics. 7th ed. California: Cengage Learning; 2010.

17. Tugwell J, Everton C, Kingma A, Oomkens DM, Pereira GA, Pimentinha DB, et al. Increasing source to image distance for AP pelvis imaging: impact on radiation dose and image quality. Radiography. 2014;20(4):351-5.

18. Williams S, Hackney L, Hogg P, Szczepura K. Breast tissue bulge and lesion visibility during stereotactic biopsy: a phantom study. Radiography. 2014;20(3):271-6.

19. Allen E, Hogg P, Ma WK, Szczepura K. Fact or fiction: an analysis of the 10 kVp 'rule' in computed radiography. Radiography. 2013;19(3):223-7.

20. Bontrager KL, editor. Tratado de técnica radiológica e base anatômica. 5a ed. Rio de Janeiro: Guanabara Koogan; 2003.

Appendix A

ESCOLA SUPERIOR DE TECNOLOGIA DA SAÚDE DE LISBOA

OPTIMAX # 14

Study: Determining exposure factors for the lowest effective dose while maintaining acceptable image quality when identifying IOFB's

Number of Image - _____

1) The Contrast between air-filled structures and the surrounding tissues / structures is...

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

2) The Trabecular patterns of the bones are visualized...

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

3) The sharpness of the orbital rim is

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

4) The visibility of the superior orbital fissure is

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

5) The amount of noise in this image is

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

6) The brightness of the visualized lesions are...

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

7) The contrast of the visualized lesions are...

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

8) The visibility of Lesion **1**:

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

9) The visibility of Lesion **2**:

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

10) The visibility of Lesion **3**:

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

11) The visibility of Lesion **4**:

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than

12) The visibility of Lesion **5**:

1 - Much worse than **2**- Worse than **3** - The same as **4** - Better than **5** - Much better than